Introduction- Explanation of the Senate's Thematic Classification System

The Senate indexes all its documentary and bibliographic collections by assigning keywords drawn from the *Senate Thesaurus*, published in 1993. This ensured comprehensive indexing of the collections, although they were not initially classified into broader fields of knowledge.

To achieve this goal without carrying out a manual, document-by-document process, it was decided to leverage the existing assignment of Thesaurus descriptors so that the contents of the collections could be automatically classified into 83 Thematic Areas (AATT), grouped into 22 broader Thematic Blocks (BBTT). In other words, the aim is to generate classification groupings—based on descriptor labelling—that automatically assign Thematic Areas and Thematic Blocks to the collections through a dedicated software solution.

The system therefore builds upon the intellectual work undertaken by the Senate's documentation services to produce a first-level grouping of Thesaurus terms into broad fields of knowledge.

We believe that this multidisciplinary and cross-cutting classification system, developed within the Senate through significant internal coordination, knowledge management and effort, may prove useful to other users of indexing systems, particularly those working with the Eurovoc Thesaurus. For this reason, we are making available—via the Open Data Portal and in .XML format—this classification of Thesaurus terms into Thematic Areas and Thematic Blocks, so that it may be freely reused and applied by anyone interested.

Structure

This document is structured as follows:

Header

Thematic blocks (22)

Block title (description)

Thematic areas (83)

Area title (description)

Subjects covered by the area (notes)

Related areas (related) (n number)

Title of related area (description)

Thesaurus terms associated with the area (term) (n number)

Term title (description)

Microthesaurus

Code and title of the microthesaurus to which it belongs

```
<blocks>
 <block num="1">
  <description>BLOCK TITLE</description>
  <areas>
   <area num="1">
    <description>AREA TITLE</description>
    <notes>List of subjects covered.</notes>
    <related>
     <related_area num="1">
      <description>Other areas with which it is related</description>
      <related_ num="n">
     </related_area>
    </related>
    <terms>
     <term num="1">
      <description>Term title</description>
      <microthesauri>
       <microthesaurus num="1">
        <description>0000 Title of the microthesaurus to which it belongs</description>
       </microthesaurus>
      </microthesauri>
     </term>
    </terms>
   </area>
  </areas>
 </block>
</blocks>
```

Some clarifications on the structure and scope of the dataset that accompanies the Senate's thematic classification system to correctly interpreting and reusing the information provided:

1. Although the *Senate Thesaurus* incorporates *EuroVoc Thesaurus* version 4.3, together with a complementary list of terms compiled by the Senate and the list of Spanish municipalities provided by the National Statistics Institute (INE), most EuroVoc terms — though not all — and a subset of the complementary list are classified within the Thematic Areas (AATT).

2. Only terms included up to *EuroVoc* version 4.3 are covered.

3. The structure of the document is as follows: *Thematic Blocks (BBTT) → Thematic Areas (AATT) → Thesaurus Terms.*

4. If you require the English translation of the Thematic Blocks (BBTT) and Thematic Areas (AATT), please contact us at sugerenciasbuscador@senado.es